UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/854,119	05/11/2001	Miroslav Trajkovic	US 010240	7390

24737 7590 04/24/2003

PHILIPS ELECTRONICS NORTH AMERICAN CORP
580 WHITE PLAINS RD
TARRYTOWN, NY 10591

EXAMINER	
AMINI, JAVID A	

ART UNIT	PAPER NUMBER
2672	

DATE MAILED: 04/24/2003

3

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 07-01)

Office Action Summary	Application No.	Applicant(s)
	09/854,119	TRAJKOVIC, MIROSLAV
	Examiner	Art Unit
	Javid A Amini	2672

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____ .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-20 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-20 is/are rejected.

7)☒ Claim(s) 1-20 is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-20 rejected under 35 U.S.C. 103(a) as being unpatentable over Woodfill, and further in view of Moezzi.

1. Claim 1.

"A method of aligning a first image to a second image, comprising: determining a first alignment approximation, based on distances between one or more points in the first image and the second image, aligning the second image to the first image, based on the first alignment approximation, to form an initially aligned second image, determining a second alignment approximation, based on distances between one or more points in the first image and the initially aligned second image, and aligning the second image to the first image, based on a combination of the first and second alignment approximation", Woodfill teaches in (col. 12, lines 57-67) the single camera may or may not be in motion . Thus, distinct images can be identified by their space/time attributes. Using a single camera, the "left" image may correspond to an image captured at one time, and the "right" image may correspond to an image captured at another time. The analysis then involves comparing successive frames; that is, if a, b, c, and d represent successive frames of images captured by the single camera, a and b are compared, then b and c, then c and d, and so

on. Similarly, the single camera may shift or move between two distinct positions (i.e., left position and right position) back and forth and the captured images are appropriately designated or assigned to either the left or right image. Woodfill does not explicitly specify the alignment based on distance between one or more points in the first image and the second image, however Moezzi teaches in (col. 28, lines 54-56) all supporting observations are used with appropriate weighting based on distance from the camera (distance between one or more points in the images) , direction of motion, and to update the position of each object. The step of second alignment based on distances is obvious because the process will continue for each frame. Woodfill teaches in (col. 32, lines 20-22) the computations will involve a combination of additions and subtractions of previously calculated arrays and image data.
Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

2. Claim 2.

"The method of claim 1, wherein aligning the second image to the first image based on the combination of the first and second alignment approximations is effected by: aligning the initially aligned second image, which is based on the first alignment approximation, to the first image, based on the second alignment approximation", Woodfill teaches in (col. 12, lines 57-67) the single camera may or may not be in motion. Thus, distinct images can be identified by their space/time attributes. Using a single camera, the "left" image may correspond to an image captured at one time, and the "right" image may correspond to an image captured at another time. The analysis then involves comparing successive frames; that is, if a, b, c, and d represent

successive frames of images captured by the single camera, a and b are compared, then b and c, then c and d, and so on. Similarly, the single camera may shift or move between two distinct positions (i.e., left position and right position) back and forth and the captured images are appropriately designated or assigned to either the left or right image.

3. Claim 3.

"The method of claim 1, wherein, determining the first alignment approximation is based on a low-resolution representation of the first and second images, and determining the second alignment approximation is based on a higher-resolution representation of the first and second images", Woodfill does not explicitly specify the higher-resolution representation of images, however Moezzi teaches in (col. 33, lines 21-26) the model shown in FIGS. 6, 7 and 8 employs an (x, y, z) world coordinate, bounding box object representation. That is, the system track object centroid and uses a bounding box to indicate presence of an object at a particular location. A voxel-based representation supports finer resolution of object shape and location.
Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

4. Claim 4.

"The method of claim 1, wherein determining at least one of the first alignment and second alignment approximations includes applying the RANSAC algorithm", the step is obvious because the RANSAC algorithms is well known in the art, however applicant should illustrates the calculations, variables and interpretation of data more in detail.

5. Claim 5.

"The method of claim 1, wherein determining the first alignment approximation includes an approximation of at least one of a rotation component and a translation component in an image space of the first and second images", Woodfill does not explicitly specify the higher-resolution representation of images, however Moezzi teaches in (col. 33, lines 21-26) the model shown in FIGS. 6, 7 and 8 employs an (x, y, z) world coordinate, bounding box object representation. That is, the system track object centroid and uses a bounding box to indicate presence of an object at a particular location. A voxel-based representation supports finer resolution of object shape and location. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

6. Claim 6.

"The method of claim 5, wherein determining the second alignment approximation includes an approximation of components of a 3x3 homographic matrix", Woodfill does not explicitly specify the 3x3 homographic matrix, however Moezzi teaches in (col. 35, lines 18-25) where M_v is the 4x4 homogeneous transformation matrix representing transformation between V and the world W. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

7. Claim 7.

"The method of claim 1, wherein determining the second alignment approximation includes an approximation of components of a 3x3 homographic matrix", Woodfill does not explicitly specify the 3x3 homographic matrix, however Moezzi teaches in (col. 35, lines 18-25) where M_v

is the 4x4 homogeneous transformation matrix representing transformation between V and the world W. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

8. Claim 8.

"The method of claim 1, wherein determining at least one of the first and second alignment approximations includes identifying corners in the first and second images based on a determination of Minimum Intensity Changes at the corners", the step is obvious because the comparison of different frame of images will provide alignment approximation to identify corners.

9. Claim 9.

"A method of tracking an object based on a first image and a second image, comprising: aligning the first and second images to form a set of aligned images, and detecting motion by comparing the set of aligned images, wherein aligning the first and second images includes: determining a first alignment approximation, based on distances between one or more points in the first image and the second image, aligning the second image to the first image, based on the first alignment approximation, to form an initially aligned second image, determining a second alignment approximation, based on distances between one or more points in the first image and the initially aligned second image, and aligning the second image to the first image, based on a combination of the first and second alignment approximations", Woodfill teaches in (col. 12, lines 57-67) the single camera may or may not be in motion . Thus, distinct images can be identified by their space/time attributes. Using a single camera, the "left" image may correspond to an image

captured at one time, and the "right" image may correspond to an image captured at another time. The analysis then involves comparing successive frames; that is, if a, b, c, and d represent successive frames of images captured by the single camera, a and b are compared, then b and c, then c and d, and so on. Similarly, the single camera may shift or move between two distinct positions (i.e., left position and right position) back and forth and the captured images are appropriately designated or assigned to either the left or right image. Woodfill does not explicitly specify the alignment based on distance between one or more points in the first image and the second image, however Moezzi teaches in (col. 28, lines 54-56) all supporting observations are used (with appropriate weighting based on distance from the camera, direction of motion, etc.) to update the position of each object. The step of second alignment based on distances is obvious because the process will continue for each frame. Woodfill teaches in (col. 32, lines 20-22) the computations will involve a combination of additions and subtractions of previously calculated arrays and image data.

Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

10. Claim 10.

"The method of claim 9, wherein determining the first alignment approximation is based on a low-resolution representation of the first and second images, and determining the second alignment approximation is based on a higher-resolution representation of the first and second images", Woodfill does not explicitly specify the higher-resolution representation of images, however Moezzi teaches in (col. 33, lines 21-26) the model shown in FIGS. 6, 7 and 8 employs

an (x, y, z) world coordinate, bounding box object representation. That is, the system track object centroid and uses a bounding box to indicate presence of an object at a particular location. A voxel-based representation supports finer resolution of object shape and location.

Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

11. Claim 11.

"The method of claim 9, further including identifying the object in the set of aligned images based on color matching", the step is obvious because the aligned images should be based on color and intensity of pixels. Woodfill teaches in (col. 12, lines 57-67) the single camera may or may not be in motion. Thus, distinct images can be identified by their space/time attributes. Using a single camera, the "left" image may correspond to an image captured at one time, and the "right" image may correspond to an image captured at another time. The analysis then involves comparing successive frames; that is, if a, b, c, and d represent successive frames of images captured by the single camera, a and b are compared, then b and c, then c and d, and so on. Similarly, the single camera may shift or move between two distinct positions (i.e., left position and right position) back and forth and the captured images are appropriately designated or assigned to either the left or right image.

12. Claim 12.

"The method of claim 9, further including determining a location of the object in each image of the set of aligned images, and determining a movement of the object by comparing the location of the object in each image", Woodfill does not explicitly specify the higher-resolution

representation of images, however Moezzi teaches in (col. 33, lines 21-26) the model shown in FIGS. 6, 7 and 8 employs an (x, y, z) world coordinate, bounding box object representation. That is, the system track object centroid and uses a bounding box to indicate presence of an object at a particular location. A voxel-based representation supports finer resolution of object shape and location. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

13. Claim 13.

"A motion detecting system comprising: a processor that is configured to: align a first image and a second image, to form a set of aligned images, by: determining a first alignment approximation, based on distances between one or more points in the first image and the second image, aligning the second image to the first image, based on the first alignment approximation, to form an initially aligned second image, determining a second alignment approximation, based on distances between one or more points in the first image and the initially aligned second image, and aligning the second image to the first image, based on a combination of the first and second alignment approximations; and compare the set of aligned images to identify motion of objects within the first and second images", Woodfill teaches in (col. 12, lines 57-67) the single camera may or may not be in motion . Thus, distinct images can be identified by their space/time attributes. Using a single camera, the "left" image may correspond to an image captured at one time, and the "right" image may correspond to an image captured at another time. The analysis then involves comparing successive frames; that is, if a, b, c, and d represent successive frames of images captured by the single camera, a and b are compared, then b and c,

then c and d, and so on. Similarly, the single camera may shift or move between two distinct positions (i.e., left position and right position) back and forth and the captured images are appropriately designated or assigned to either the left or right image. Woodfill does not explicitly specify the alignment based on distance between one or more points in the first image and the second image, however Moezzi teaches in (col. 28, lines 54-56) all supporting observations are used (with appropriate weighting based on distance from the camera, direction of motion, etc.) to update the position of each object. The step of second alignment based on distances is obvious because the process will continue for each frame. Woodfill teaches in (col. 32, lines 20-22) the computations will involve a combination of additions and subtractions of previously calculated arrays and image data.

Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

14. Claim 14.

"The motion detecting system of claim 13, wherein the processor is configured to: determine the first alignment approximation by processing a low-resolution representation of at least one of the first and second images, and determine the second alignment approximation by processing a higher-resolution representation of the first and second images", Woodfill does not explicitly specify the higher-resolution representation of images, however Moezzi teaches in (col. 33, lines 21-26) the model shown in FIGS. 6, 7 and 8 employs an (x, y, z) world coordinate, bounding box object representation. That is, the system track object centroid and uses a bounding box to

indicate presence of an object at a particular location. A voxel-based representation supports finer resolution of object shape and location.

Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

15. Claim 15.

"The motion detecting system of claim 13, further including one or more cameras for producing the first and second images", Woodfill illustrates in Fig. 1.

16. Claim 16.

"The motion detecting system of claim 13, further including a memory for storing a representation of a target image, and wherein the processor is further configured to identify a target within the set of aligned images, based on the representation of the target image", Woodfill illustrates in Fig. 65. And also Woodfill teaches in (col. 8, lines 62-67) the hardware aspect of the present invention represents a parallel pipelined computing system designed to perform data set comparisons efficiently and at low cost. Data is processed in a systolic nature through the pipeline. This image processing system provides high performance via high computational density, high memory bandwidth, and high I/O bandwidth.

17. Claim 17.

"The motion detecting system of claim 16, wherein the representation of the target image is a characterization based on color content of the target image", the step is obvious because the aligned images should be based on color and intensity of pixels. Woodfill teaches in (col. 12, lines 57-67) the single camera may or may not be in motion. Thus, distinct images can be

identified by their space/time attributes. Using a single camera, the "left" image may correspond to an image captured at one time, and the "right" image may correspond to an image captured at another time. The analysis then involves comparing successive frames; that is, if a, b, c, and d represent successive frames of images captured by the single camera, a and b are compared, then b and c, then c and d, and so on. Similarly, the single camera may shift or move between two distinct positions (i.e., left position and right position) back and forth and the captured images are appropriately designated or assigned to either the left or right image.

18. Claim 18.

"The motion detecting system of claim 13, further including determining a location of an object in each image of the set of aligned images, and determining a movement of the object by comparing the location of the object in each image", Woodfill teaches in (col. 12, lines 57-67) the single camera may or may not be in motion. Thus, distinct images can be identified by their space/time attributes. Using a single camera, the "left" image may correspond to an image captured at one time, and the "right" image may correspond to an image captured at another time. The analysis then involves comparing successive frames; that is, if a, b, c, and d represent successive frames of images captured by the single camera, a and b are compared, then b and c, then c and d, and so on. Similarly, the single camera may shift or move between two distinct positions (i.e., left position and right position) back and forth and the captured images are appropriately designated or assigned to either the left or right image.

19. Claim 19.

"The motion detecting system of claim 13, wherein determining the first alignment approximation includes an approximation of at least one of a rotation component and a

translation component", Woodfill teaches in (col. 12, lines 57-67) the single camera may or may not be in motion. Thus, distinct images can be identified by their space/time attributes. Using a single camera, the "left" image may correspond to an image captured at one time, and the "right" image may correspond to an image captured at another time. The analysis then involves comparing successive frames; that is, if a, b, c, and d represent successive frames of images captured by the single camera, a and b are compared, then b and c, then c and d, and so on. Similarly, the single camera may shift or move between two distinct positions (i.e., left position and right position) back and forth and the captured images are appropriately designated or assigned to either the left or right image.

20. Claim 20.

"The motion detecting system of claim 19, wherein determining the second alignment approximation includes an approximation of components of a 3x3 homographic matrix", Woodfill does not explicitly specify the 3x3 homographic matrix, however Moezzi teaches in (col. 35, lines 18-25) where M_v is the 4x4 homogeneous transformation matrix representing transformation between V and the world W. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate the teaching of Moezzi into Woodfill in order to improve the lack of computation power.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Javid A Amini whose telephone number is 703-605-4248. The examiner can normally be reached on 8-4pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Michael Razavi can be reached on 703-305-4713. The fax phone numbers for the organization where this application or proceeding is assigned are 703-746-8705 for regular communications and 703-746-8705 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is 703-306-0377.

Javid A Amini
Examiner
Art Unit 2672

Javid Amini
April 17, 2003

MICHAEL RAZAVI
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600